

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 04:04:14 ; Search time 6128.19 Seconds  
(without alignments)  
2020.032 Million cell updates/sec

Title: US-09-526-329-38

Perfect score: 1152  
Sequence: 1 gaggcgggtccggttgcag.....aaaaaaaaaaaaaaaa 1152

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapept 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpl:\*  
6: em\_estlba:\*  
7: em\_estlro:\*  
8: em\_estrov:\*  
9: em\_hic:\*  
10: gp\_est1:\*  
11: gp\_est2:\*  
12: gp\_hic:\*  
13: gp\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rnd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	664.4	57.7	674	AM978572	EST390681
2	606.6	52.7	657	AM370369	RC1-BT025
3	498.2	43.2	508	AA825207	OC66607.S
4	493	42.8	495	AM575965	UI-HF-B10
5	454.4	39.4	456	AM1379831	ts29409.X
6	407	35.3	407	AM1224367	q405107.X
7	396.4	34.4	398	AM292092	UI-H-B12-
8	374.4	32.5	376	AM1832007	w199402.X
9	372.8	32.4	525	AM1250843	qx10404.X
10	349.4	30.3	358	AM404913	UI-HF-B10
11	331.4	28.8	338	AA825205	OC66605.S
12	312.8	27.2	328	AM003249	wg64e08.X

C 13	304	26.4	315	10	AI434322	AI434322 t141d01.x
C 14	283.4	24.6	300	10	AI250350	AI250350 GS04C01.X
C 15	262.8	22.8	492	13	AO589625	HS_2132.A
C 16	261.8	22.7	265	10	AM768336	AM768336 HK72905.X
C 17	239.2	20.8	463	13	AO572349	AO572349 HS_2137.B
C 18	189	16.4	389	10	AM480639	AM480639 32244.MAR
C 19	173.4	15.1	560	10	BE685287	BE685287 187456.MAR
C 20	169.4	14.7	457	13	AO518451	AO518451 HS_5101.A
C 21	166	14.4	178	10	AI468861	AI468861 t143b01.x
C 22	143.2	12.4	148	10	AA689513	AA689513 ns17g04.X
C 23	125.8	10.9	161	11	BF082894	BF082894 CV1-BT082
C 24	119	10.3	282	10	AM345961	AM345961 26836.MAR
C 25	93.8	8.1	472	11	BF605796	BF605796 27232.MAR
C 26	73	6.3	365	10	AM014278	AM014278 UI-H-B10
C 27	63.6	5.5	537	10	AI391250	AI391250 mb58c01.Y
C 28	62.8	5.5	431	10	AM405686	AM405686 UI-HF-B10
C 29	62.8	5.5	449	11	BF128815	BF128815 601811195
C 30	62.8	5.5	493	10	BE267581	BE267581 601124363
C 31	62.8	5.5	494	10	AM500764	AM500764 UI-HF-BN0
C 32	62.8	5.5	494	10	AM732980	AM732980 db18d04.Y
C 33	62.8	5.5	511	10	BE398017	BE398017 601290574
C 34	62.8	5.5	515	10	BE267967	BE267967 601125372
C 35	62.8	5.5	535	10	BE268990	BE268990 601184424
C 36	62.8	5.5	547	10	BE269500	BE269500 601184704
C 37	62.8	5.5	553	10	BE513898	BE513898 601316486
C 38	62.8	5.5	554	11	BI225045	BI225045 602949269
C 39	62.8	5.5	554	11	BI225862	BI225862 602950339
C 40	62.8	5.5	559	11	BI261356	BI261356 602953358
C 41	62.8	5.5	566	10	BE294557	BE294557 601174505
C 42	62.8	5.5	570	11	BI226934	BI226934 602952063
C 43	62.8	5.5	574	10	AM732867	AM732867 db10h10.Y
C 44	62.8	5.5	576	10	AM328037	AM328037 dt03g06.X
C 45	62.8	5.5	578	11	BG759756	BG759756 602711155

#### ALIGNMENTS

RESULT 1  
AM978572/c  
LOCUS AM978572 674 bp mRNA  
DEFINITION EST390681 MAGE resequences, MAGP Homo sapiens CDNA, mRNA sequence.  
ACCESSION AM978572.1  
VERSION AM978572.1 GI:8169840  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 674)  
Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt  
,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
Quackenbush,J.  
Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
unpublished (2000)  
JOURNAL Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnget@igf.org  
Plate: 394  
Seq primer: Forward.  
FEATURES  
Source 1. 674  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGP"  
/note="Vector: pBluescriptKm"  
BASE COUNT 185 a 178 c 191 g 120 t  
ORIGIN

Query Match 57.7%, Score 664.4; DB 10; Length 674;  
 Best Local Similarity 99.1%; Pred. No. 1.1e-11;  
 Matches 668; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 453 ggcctggtgtctctcctcagcccccctcagtgagatcttcagtcctcgtctctgttg 512  
 Db 674 ggctggtgtgtctcctcagcccccctcagtgagatcttcagtcctcgtctctgttg 615  
 OY 513 cacaccagatagatcctcctctcagagagaagagagagagagagagagagagagag 572  
 Db 614 cacaccagatagatcctcctcctcagagagaagagagagagagagagagagagag 555  
 OY 573 tctcgtctctcctcagatcagtcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 632  
 Db 554 tctcgtctctcctcagatcagtcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 495  
 OY 633 tggcctggtatgggtctcactcagtgaggccctcgtcgtcgtcgtcgtcgtcgtcgtc 692  
 Db 494 tggcctggtatgggtctcactcagtgaggccctcgtcgtcgtcgtcgtcgtcgtcgtc 435  
 OY 693 tgcagctgtggtgtctctcctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 752  
 Db 434 tgcagctgtggtgtctctcctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 375  
 OY 753 caatccctcctcctacccacagagatgtgtccagcagagcctcagacccctcagtcgc 812  
 Db 374 caatccctcctcctacccacagagatgtgtccagcagagcctcagacccctcagtcgc 315  
 OY 813 agcccgagtcagtaaacctcaccacagcagcagcagcagcagcagcagcagcagcagc 872  
 Db 314 agcccgagtcagtaaacctcaccacagcagcagcagcagcagcagcagcagcagcagc 255  
 OY 873 agggagagccagctctcgtatgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 932  
 Db 254 agggagagccagctctcgtatgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 195  
 OY 933 tggcccgaaatagatcactcagtcagcaagcaatgataccagcagcagcagcagcagc 992  
 Db 194 tggcccgaaatagatcactcagtcagcaagcaatgataccagcagcagcagcagcagc 135  
 OY 993 ctgcacagctggagagatcagagagagagagagagagagagagagagagagagagag 1052  
 Db 134 ttgcacagctggagagatcagagagagagagagagagagagagagagagagagagag 75  
 OY 1053 cagctcag 1112  
 Db 74 cagctcag 15  
 OY 1113 ttattgtcactgt 1126  
 Db 14 ttattgtcactgt 1

RESULT 2  
 AM370369/c 657 bp mRNA EST 04-FEB-2000  
 LOCUS RCI-BT0255-181099-012-f04 BT0255 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION  
 ACCESSION AM370369  
 VERSION AM370369.1 GI:6875023  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 657)  
 AUTHORS HGP http://www.ludwig.org.br/ORESTES.  
 TITLE The FAPESP/LICR Human Cancer Genome Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP.

Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL:  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC1&l2=RC1-BT0255-  
 181099-012-f04&l3=1999-10-18&l4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 29  
 High quality sequence stop: 657.  
 Location/Qualifiers  
 1..657  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BT0255"  
 /dev\_stage="Adult"  
 /note="Organ: Breast; Vector: puc18; Site\_1: Sma1; Site\_2:  
 Sma1; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No. 196  
 716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

BASE COUNT 174 a 174 c 187 g 122 t  
 ORIGIN

Query Match 52.7%; Score 606.6; DB 10; Length 657;  
 Best Local Similarity 97.0%; Pred. No. 3.9e-101;  
 Matches 618; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 485 gatcttcagtcacccctcctcgtctgttgacacacccagatagcctcttcgagagagag 544  
 Db 657 gatcttcagtcacccctcctcgtctgttgacacacccagatagcctcttcgagagagag 598  
 OY 545 gcaatagagccctcagacactcagttcctcgtctcgtctcgtctcgtctcgtctcgtc 604  
 Db 597 gcaatagagccctcagacactcagttcctcgtctcgtctcgtctcgtctcgtctcgtc 538  
 OY 605 tccacactcagctcgtcacttagagagagcagcagcagcagcagcagcagcagcagcagc 664  
 Db 537 tccacactcagctcgtcacttagagagagcagcagcagcagcagcagcagcagcagcagc 478  
 OY 665 gtgtcgtcagcaggttccctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 724  
 Db 477 gtgtcgtcagcaggttccctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 418  
 OY 725 ctcatgtcgtcctcctcagatgcacccccaatccctcctacacacagagatgtcgc 784  
 Db 417 ctcatgtcgtcctcctcagatgcacccccaatccctcctacacacagagatgtcgc 358  
 OY 785 ccagcagagcctccagacccccaagtgagcagctcgtgattggaactcaccatcggcaggc 844  
 Db 357 ccagcagagcctccagacccccaagtgagcagctcgtgattggaactcaccatcggcaggc 298  
 OY 845 agtggtcgtttagagatgagcattagagagagagagagagagagagagagagagagagag 904  
 Db 297 agtggtcgtttagagatgagcattagagagagagagagagagagagagagagagagagag 238  
 OY 905 ccgtggtgatacagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 964  
 Db 237 ccgtggtgatacagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 178  
 OY 965 atgtaacccggagcctcagtagagcaccatcagcagagtgagagagagagagagagagag 1024  
 Db 177 atgtaacccggagcctcagtagagcaccatcagcagagtgagagagagagagagagagag 118  
 OY 1025 ggcctggtcctcgtctatcagtcctcctcagctcagcagagagagagagagagagag 1084  
 Db 117 ggcctggtcctcgtctatcagtcctcctcagctcagcagagagagagagagagagag 58  
 OY 1085 tgacctcagctcgtgtaatacttattgtc 1121







Db	278	CATTACAGGGAGCCCAAGCTCGATGTGGCACTTGGATGCCCTGTGGATATCAATTTCGCTG	219
Qy	927	aaactctggccggcaataatgatatacaagtcgctggagcagcaatctatcacccggagactcaatgta	986
Db	218	AACATCTTGGGCCCAATATGATATCAAGTCTGGACCAAGCAATATACACGGAGACCTCAGTA	159
Qy	987	gcccatctgacagatggaggagcattggaaggaatggatcttgacctgtgctcttgattca	1046
Db	158	GCCCATCTTGCACAGTGGGGACATGGAAGGAGGGTTTGGCCCTGATCTTCGCTATTCA	99
Qy	1047	gtccttaagctcaaggaaaggatgctatgctccgtgaagtgagacctcaagtaactggttaat	1106
Db	98	GTCCCTTCACGCTACCGAAGGGATGCTAGTCCGTAAAGTGACTCTCAGTACTGTTAAAT	39
Qy	1107	taaaccttatctgctcaatgctcaaaaaaataaaaaa	1144
Db	38	TAAACTTATTGCTCATCTGTCAAAAAAAAAAAAAAAAAA	1

[illegible]

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (Pages 1 to 376)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapdb-rt@mail.nih.gov](mailto:cgapdb-rt@mail.nih.gov)

The **Sequences** tab contains the Washington University Genome Sequencing Center  
 clone distribution: NCBI-CCAP clone distribution  
 www.bio.1ml.gov/hbrp/image/image.html  
 Insert length: 1100 Std Error: 0.00  
 Seq primer: -40MP from Gibco  
 High quality sequence stop: 375.  
 Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2410922"
/clone_id="NCI_CGAP_Lym12"
/tissue_type="Lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/notes="Organ: Lymph node; Vector: pCMV-SPORT6; Site_1:
SalI; Site_2: NotI; Cloned unidirectionally. Primer:
019g dt. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"

```

	Query Match	32.5%	Score 374.4	DB 10	Length 376
	Best Local Similarity	99.7%	Pred. No. 9,7e-59		
	Matches 375	Conservative	0	Mismatches 1	Indels 0
				Gaps	0
Oy	753 caaatcccttcataccacccaggatgctggccagccagcctccaaacccccagtc	812			
Db	376 CAAATCCCTTCTATACCCACCGAGATGTGTGCCACGCCAGCCTCCAGACCCCGATGC	317			
Oy	813 agctcgtgattgaaactccacatccagtcagcagcagtggttcgttttaagaatgtgcattg	872			

Db	316	AGCGCGGATTGTGGAACCTACCCATCGCGAGCGAGTGGTTCGGTTTAAGATGAGATGGCATTAG	257
OY	873	aggagagcccaagctcgtggaatgagacttgatgccctctggatgatcaagttcgtgcacact	932
Db	256	AGGAGCGCCAGATCGTGAGATGTGGACTGTGATGCCCTGTGGGTATATAGATTGCTGCACACTT	197
OY	933	tgagcccaaatatgatcagctgctcgagcaagcaatctacaccgagacctcaagcgagcccat	992
Db	196	TGGCCGGAATATGATCCAGGCGCGAGCAGACAGCATTATACCGGAGCCCTCAGTAGACCCAT	137
OY	993	ctgcaacagctgagggagacatagagaggaatggatgttagcctgtagctatgtctatcagtcct	1052
Db	136	CTGCACAGTGGGGAGCATGGAGGATGGGCTTTGGCCTGTCCTTGCTTATTCACATCCTT	77
OY	1053	caagctaaaggaaggaatgctagtcgttgaagctgactacacagacagctgattataaact	1112
Db	76	CAGCTTACCGAAGAGATGCTAGTCCCTGATAGAGTACTCAGACAGTACGTGTTAATTAACT	17
OY	1113	ttatctgctcaactgtca	1128
Db	16	TTAATGCTCAGCTGA	1

RESULT	9
A1250843/C	
LOCUS	
DEFINITION	525 bp mRNA
ACCESSION	gq10404.x1 NC1_CGAP_Lym12 Homo sapiens CDNA clone
VERSION	A1250843.1 GI:3847372
EST	21-DEC-1998
IMAGE:	2000510 3'

ORGANISM	REFERENCE AUTHORS TITLE	JOURNAL COMMENT
Homo sapiens		
Akaryotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
1 (bases 1 to 525)		
NCI-CCAP <a href="http://www.ncbi.nlm.nih.gov/ncicap">http://www.ncbi.nlm.nih.gov/ncicap</a> .		
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
Tumor Gene Index		
Unpublished (1997)		
Contact: Robert Strausberg, Ph.D.		
ncicg@nci.nih.gov		
ncicg@nci.nih.gov		

```

UNKNOWN library type
Insert length: 1149      Std Error: 0.00
Seq primer: -400P from GIBCO
High quality sequence stop: 336.
Location/Qualifiers
1. .525
FEATURES
source
```

```

/db_xref="taxon:3606"
/clone="IMAGE:2000910"
/clone_1lb="NCI CGAP Lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pGMV-SPOK6; Site_1:
Salt; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.25 kb. llyf technologies
catalog #: 11547-015"
BASE COUNT      143 a      141 c      136 g      104 t      1 others
ORIGIN

```

Query Match	32.48	Score 372.8	DB 10	Length 525
Best Local Similarity	99.28	Pred. No. 1,7e-58		
Matches 374	Conservative	0	Mismatches 5	Indels 0
			Gaps 0	
Qy 752	ccaatccccccttaaccacacagatgtgtgcccagccagagccctcaacgaccccaagtg	811		
Db 377	CAAAATCCCTTCATACCACACAGNATGTGTGCCACACCAGGCTCTCAGACCCCAAGT	318		
Qy 812	cagctcgttgattggaaactaccatctggcagagcggtgttcggtttaaagatgacatga	871		

QY	48	atggagcctccgaagcctctctgtgctatgaggggtgcacctggccgcgtctgttgatccaaagg	87
Db	8	ATGGCCTCCGAACCTTCTGTGGCTGAGGGGTCCCGCCGCGTGTGGATCCAAAGG	67
QY	88	ccctgcatctacgaagatctagagagggagagacctggtgactgtgtctgtgcgttcaat	147
Db	68	CCTGGCATCTACGAAGATATAGAGAGGGAGAACCTGGAGACTGTGGTCTGGGGTTCAAT	127
QY	148	cccttcgtaaggaatctggtccaaaggcctcccaaggccaagatctgaaccacgaatca	207
Db	128	CCCTTGGCTAGGGGAATGGGSCAGGGCCTCCACAGGCACAGATATGMACCAGCATCCA	187
QY	208	gtgcacattgtggcaagatctgagatctatccacggggagctatctctcccgccagatgcc	267
Db	188	GTGCACATTGTGGCAGATGSCAMTGCATACCCGGAGACATCTCTCTCCGGCCAGATGCC	247
QY	268	ttcttccaaagctgcccgcgtgtgtgcagctcttaccgccggaggaagtaacgaagcaagat	327
Db	248	TTCTTCCAAAGCTGCCCGCGTGGCAGACTCTACCCGGGAGGAATACCGACAGCGGAT	307
QY	328	tccagattctcggaaatagagacacatcggccagatctgactctatgtagagag	378
Db	308	TCCAGATTCTCGGAATATACACCAATGGCCAGATTAAGTACTTATAGAGAGC	358

RESULT	11
AA825205/c	
LOCUS	
DEFINITION	AA825205 38 bp mRNA EST 21-APR-1998
ACCESSION	0606605.s1 NC1_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1354668 3'
VERSION	mRNA sequence.
KEYWORDS	AA825205
FEATURES	AA825205.1 GI:2898502
FASTA	

**ORGANISMS**  
Eukaryotes: Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Placentalia; Primates; Carnivori; Homnidae; Homo.  
1 (bases 1 to 338)  
[NCI-GAP http://www.ncbi.nlm.nih.gov/ncicgap](http://www.ncbi.nlm.nih.gov/ncicgap),  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

**JOURNAL**  
Contact: Robert Strausberg, Ph.D.

**COMMENT**

FEATURES  
source  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrived by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/URL at:  
[www-bio.ln1.gov/bhrp/image/image.html](http://www-bio.ln1.gov/bhrp/image/image.html)  
Insert length: 1077 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 135.  
Location/Qualifiers  
1. 338

FEATURES

SOURCE

poly(1,4-butylene terephthalate) (PBT) with a modified linker. Site 1: Not I; Site 2: Eco RI; 1st strand was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20<sup>+</sup>, IgD<sup>+</sup>), provided by Dr. Louis M. Staudt (NCI). Dr. David Allman (NCI) and Dr. Gerald Marti (CBER) cDNA synthesis was primed with a Not I - oligo(dT) primer: 5'-TGTTGACCATGTGAAGTGAGGAGGAGCCGTCATTTTTTTTTT-3'. Double-stranded cDNA was ligated to Eco RI adaptors



(pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 80 a 91 c 69 g 98 t  
ORIGIN

Query Match 28.8%; Score 331.4; DB 10; Length 338;  
Best Local Similarity 99.7%; Pred. No. 6.9e-51;  
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 820 gattggaactacacatcgagcagcagtggttcggttaagagatgacgaaggagc 879  
DB 338 GATTGGAAACTCACCATCGGACGAGTGGTTGCTTAAAGATGCGATTACGAGGAGC 279  
OY 880 ccagtcgagtgactgactgagccctggtgatatcagcttcgcgaacttggccg 939  
DB 278 CCACTCTGAGTGTGACTTGATGCCCTGTGGTATCAGTTCTGCGACACTTTGGCCG 219  
OY 940 aaataaaccagtcgtgagcaagcaatgtacacggagcctcagtgagccatctgaca 999  
DB 218 AAATAATCATCAGTCTGAGCAACCAATGTACCGGAGCCTCAGTAGCCCATCTGCACA 159  
OY 1000 gtcggaagcagtgaggaagtggttgccctgtgcttctgcttcaatcagctcagctca 1059  
DB 158 GTGGGAGCATGAGAGGATGGGTTGGCTGTGCTCTATTCAGTCACTCAGCTCA 99  
OY 1060 cggaaagagtgatgctgctgagagtgagctcagctcagctcagctcagctcagctc 1119  
DB 98 CGGAGAGGATGCTAGTCCGTGAGAGTGACCTCAGTCACTGCTTAACTTATTCG 39  
OY 1120 tcactgtcaaaaaaataaaaaaataaaaaa 1152  
DB 38 TCACTGTCAAAAAAATAAAAAAATAAAAAA 6

RESULT 12  
AM003249 328 bp mRNA EST 08-MAR-2000  
LOCUS w964e08.x1 NCI\_CGAP\_GC6 Homo sapiens cDNA clone IMAGE:2476070 3'  
DEFINITION mRNA sequence.

ACCESSION AM003249.1 GI:5850165  
VERSION EST.  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 328)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdrrp/image/image.html](http://www-bio.llnl.gov/bdrrp/image/image.html)  
Insert Length: 384 Std Error: 0.00  
Seq primer: -400P from Gibco.

FEATURES  
source  
1..328  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2476070"  
/clone\_11b="NCI\_CGAP\_GC6"

/tissue\_type="pooled germ cell tumors"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI\_CGAP\_GC6 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 84 a 95 c 73 g 76 t  
ORIGIN

Query Match 27.2%; Score 312.8; DB 10; Length 328;  
Best Local Similarity 99.1%; Pred. No. 1.7e-47;  
Matches 325; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 807 cagtcagctcgtgattggaactcacatcgagcagcagtggttcg9-11taagagaag 865  
DB 328 CAGTCGAGCTGCTGATTGGAACTCAGCATCGGAGGAGTGGTTGCTTTAAAGANG 269  
OY 866 gcattagaggaagccacagctctgagtgacttgatgccctgttgatcagtttgt 925  
DB 268 GCATTAGAGGAGGCCAGTCTGAGTGTGACTTGGATGCCCTGTGGTATCAGTTTCT 209  
OY 926 gaacatttgcccgaaatagatccagctgctgagcaagcaatgtacaccgagccatg 985  
DB 208 GACACTTGGCCCCAATAAGATCAGTCTGAGCAACGATGTACACCGAGCCTCAGTG 149  
OY 986 agccatctcgaagagtgaggaagatgagtggttgccctgttcttcttcttcttc 1045  
DB 148 AGCCACTCTGCACTGAGTGGAGCATGAGATGGGTTGGCTTCTTCTGCTTATTC 89  
OY 1046 agtccttcagctcagcaggaagtgatgctgctgagtgagtgacccagctggtta 1105  
DB 88 AGTCCTTCAGCTCAGCAGGAGGATGCTGCTGAGGAGTGACCTCCAGTACGATGTTAA 29  
OY 1106 ttaacttatgtctcactgtcaaaaaa 1133  
DB 28 TTAACCTTATTTGCTCACTGTCAAAAAA 1

RESULT 13  
A1434322 315 bp mRNA EST 30-MAR-1999  
LOCUS t141d01.x1 NCI\_CGAP\_Lym12 Homo sapiens cDNA clone IMAGE:2133025 3'  
DEFINITION mRNA sequence.

ACCESSION A1434322  
VERSION A1434322  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 315)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Life Technologies catalog #: 11547-015  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdrrp/image/image.html](http://www-bio.llnl.gov/bdrrp/image/image.html)  
Insert Length: 1751 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 308.

FEATURES  
source  
1..315  
Location/Qualifiers



```

/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10s"
/note="Organ: lymph node; Vector: PCMV-SPORE6; site_1:
Sali; Site_2: Noli; Cloned unidirectionally. Primer:
Oligo dt. Average insert size 1.25 Kb. Life Technologies
catalog no. 11547-015"
BASE COUNT      82 a      88 c      64 g      66 t
ORIGIN
Query Match      24.6%;   Score 283.4;   DB 10;   Length 300;

```

	Query Match	Local Similarity	24.6%	Score 283.4	DB 10	Length 300
	Best Match	Similarity 19.6%	Pred. No. 4e-42			
	Matches 284	Conservative 0	Mismatches 1	Indels 0	Gaps 0	
QY	845	aaagagttcaggttaaagaagatgcgttaagaagaagccacgctcgtgatgtgagacttgaatgc	904			
DB	300	AGAGGTTCGGTTTAAAGATGGCATTTAGAGAGAGCCAGCTGTGATGTGAGCATTTGATATC	241			
QY	905	cccttgaggatcaagcttgcgtgcgaactctggcccgaaatagatccagtgcgtgcgaagca	964			
DB	240	CCTGGGGTATCACTTCTCTCTGACACTCTTGGCCGAATATGATTCAGTCTGGAGCAAGA	181			
QY	965	atgtacacaggaagcctcaagtgaagccactctgcacagctgaggagacatgagagagatgagttt	1024			
DB	180	ATGTACACCGGAGCCCTCAGTAGGCCATCTGCACAGTGGGGAGCATGTGAGGATGGGGTTT	121			
QY	1025	ggcctgtgcttctgcatttcagtccttcaagctcaaggaagagatgcatgctcgttaag	1084			
DB	120	GGCTGTGCTTCTGCTTATTCAGTCCTTCAGTCAACGGAAGGATGCTAATCCGTGAAG	61			
QY	1085	tgacctacagtaactggtatlaataacttiatgtctcaactgcca	1129			
DB	60	TGACTCACAGTACGTGTATTAATTAACCTTATTTCTCAGCTCCA	16			

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AO589625											
AO589625	492 bp DNA	AO589625	HS_2132.AL.D12.T7C	CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2132 COL-23 Row=G, DNA sequence.	AO589625	GI:5020678	GSS	08-JUN-1999			
						human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
							1 (bases 1 to 492)	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,			
								Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,			
								Reider,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and			
								Hood,L.			
									Sequence-tagged connectors: A sequence approach to mapping and		
									scanning the human genome		
									Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)		
									Contact: Mahairas GG, Wallace JC, Hood L		
									99380589		

```

FEATURES
  source
    Location/Qualifiers
      1..492
        /organism="Homo sapiens"
        /db_xref="taxon:9606"

```

